

---

# Integrated Performance Analysis of Distributed Computer Systems

---



**Fraunhofer** Institut  
Techno- und  
Wirtschaftsmathematik



---

## Integrated Performance Analysis of Distributed Computer Systems

Dirk Merten   Fraunhofer ITWM

Kaiserslautern, Germany

Workshop on Performance Characterization, Modeling and  
Benchmarking for HPC Systems

May 2003, Emeryville

---



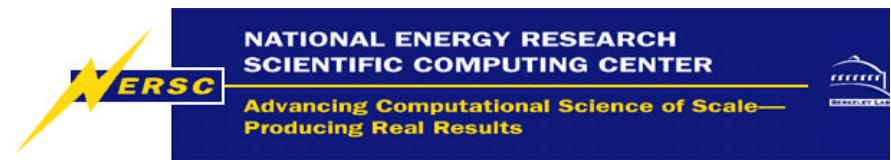
Bundesministerium  
für Bildung  
und Forschung



Fraunhofer  
Institut  
Techno- und  
Wirtschaftsmathematik



...T...Systems



## Aims:

- Development and distribution of scalable, portable and realistic benchmarks for todays and future parallel computers
- Low Level Benchmarks to characterize computer architectures
- Benchmarks of commercial software
- Performance prediction
- make benchmarking easier

---

Funding: 1.5 Mio \$, 3 years, 20 men years

Fraunhofer  
Institut  
Techno- und  
Wirtschaftsmathematik

IPACS  
Benchmark

---

## Low-level Benchmarks

---

## Low-level Benchmarks

Message Passing Benchmark PMB

[www.pallas.com/e/products/pmb](http://www.pallas.com/e/products/pmb)

Cache Benchmark CACHEBENCH  
(Philip Mucci, Univ. Knoxville)

[icl.cs.utk.edu/projects/lrcbench](http://icl.cs.utk.edu/projects/lrcbench)

IO Benchmark IOZONE (W. Norcott, D. Capps)

[www.iozone.org](http://www.iozone.org)

Try to characterize performance by few numbers

---

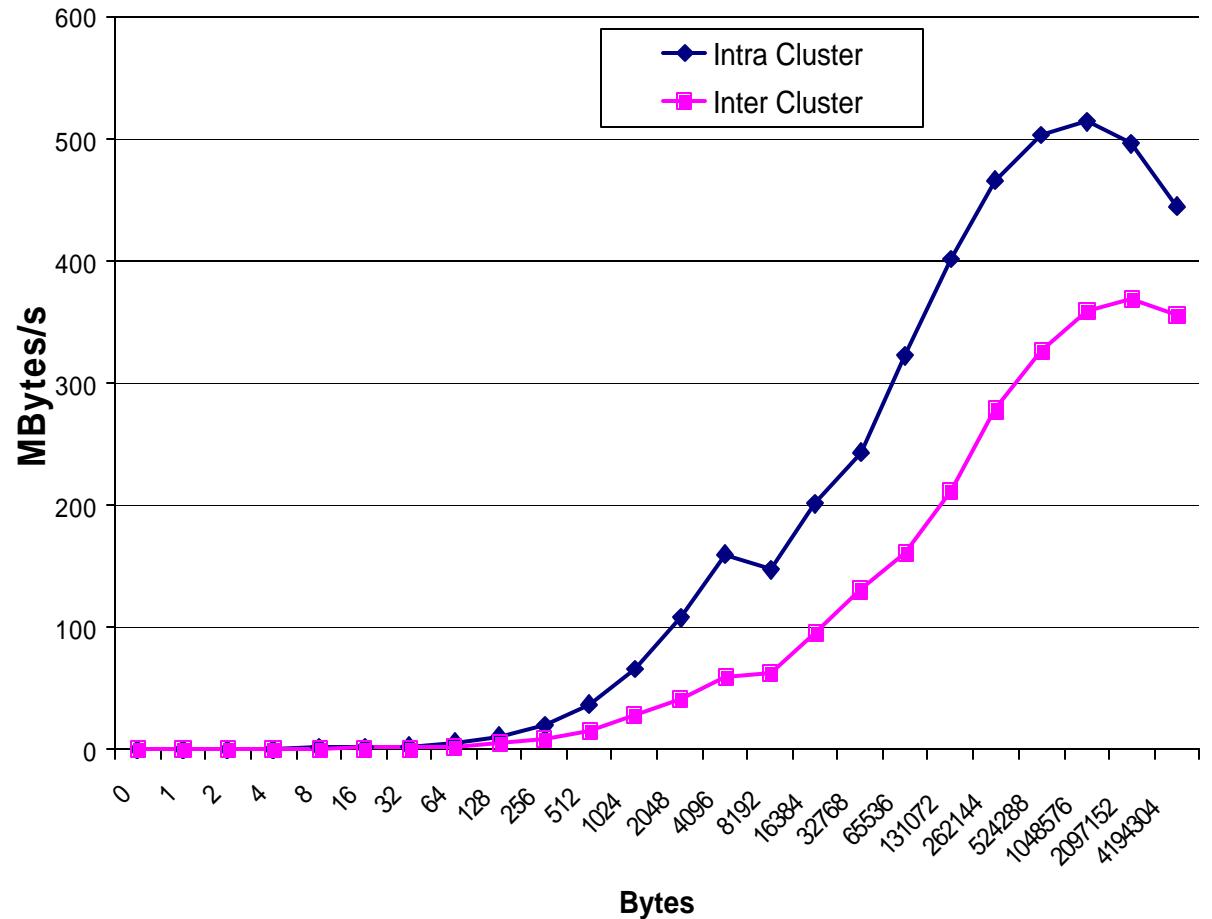
# PMB Example

Pallas MPI Benchmark

Checks many MPI communication patterns

Automatically detects clustering and reports intra-cluster and inter-cluster performance

## IPACS Message Passing Benchmark (IBM SP3)



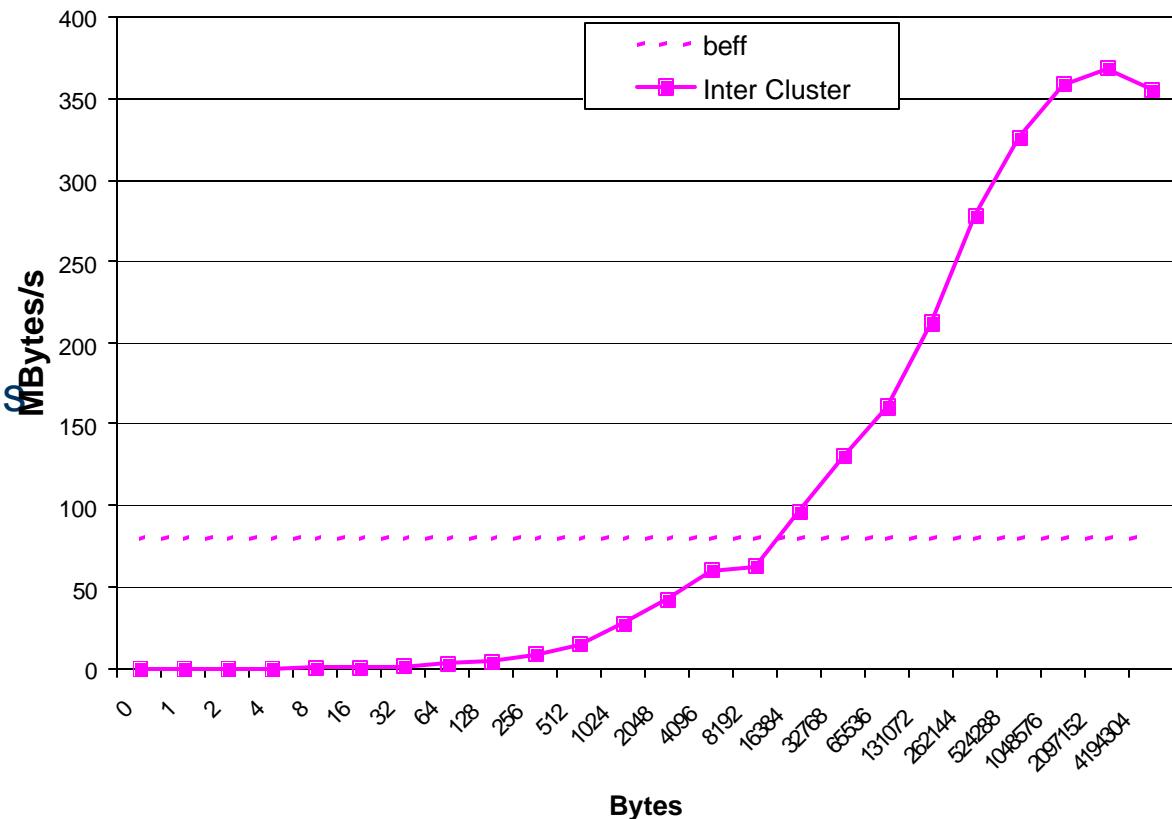
# Effective Bandwidth

MPI Communication performance as single number

N/2-N/2 ping-pong, integration over message-lengths, average over various patterns

Can be compared to CPU performance to compute „balance“

IPACS Message Passing Benchmark (IBM SP3)



# B\_EFF results

	Myrinet/Linux	Quadrix/Linux	IBM SP3
	LANL	LLNL	NERSC
2	53	87	
4	53	87	
8	51	84	
16	48	82	1node 39
32	45	78	2nodes 30
64	42	76	
80			5 nodes 21
128	39	72	
160			10 nodes 18
256	34	64	
320			20 nodes 16
384	33	60	
512		58	32 nodes 16
768		48	



# CACHEBENCH Example

Itanium 733 MHz Redhat Linux 7.2

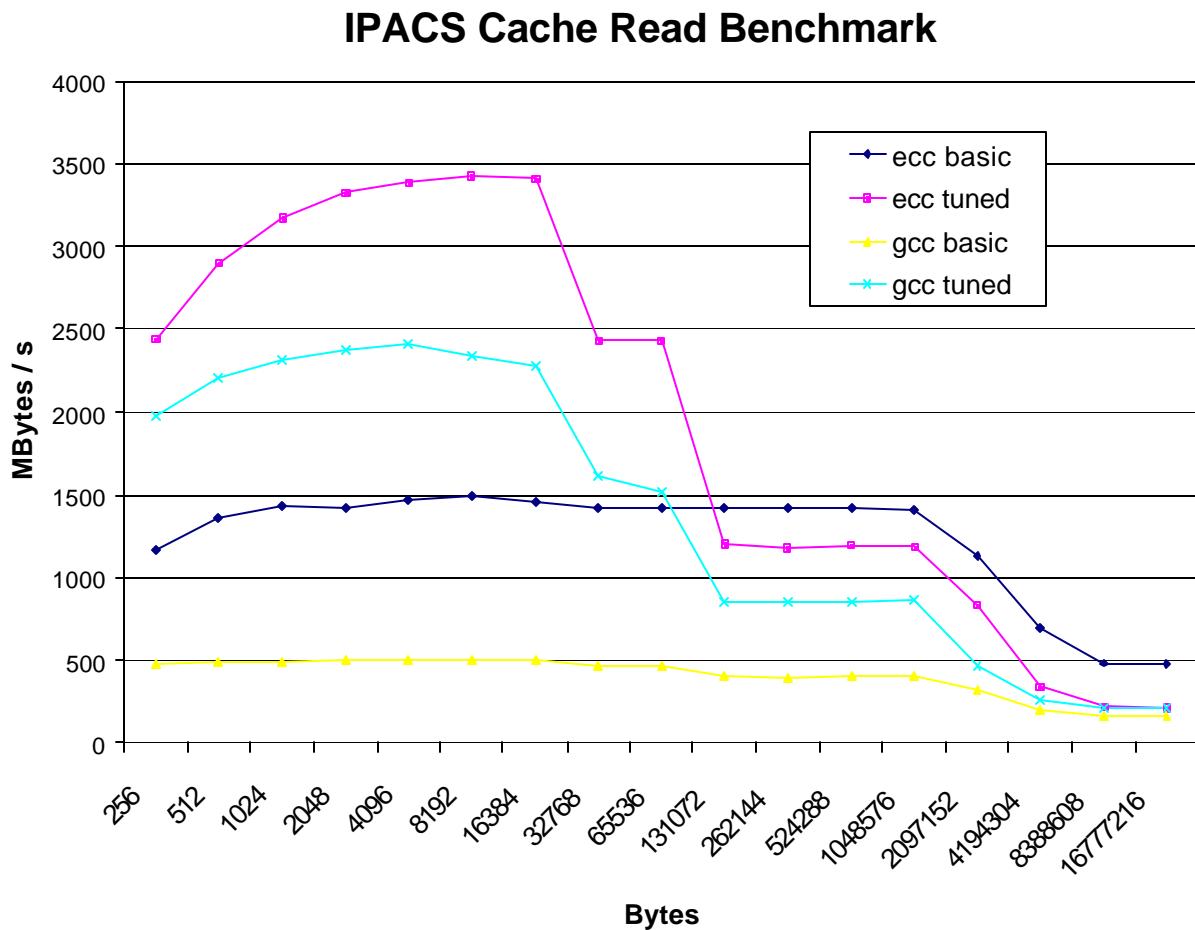
Caches: L1 16 KB, L2 96 KB

Diversity of operations:

read, write, modify  
from/to memory to/from register  
variable

hand-tuned versions of all (8 fold  
unrolling)

memset, memcpy



# IOZONE Example

Itanium 733 MHz Redhat Linux 7.2

Caches: L1 16 KB, L2 96 KB

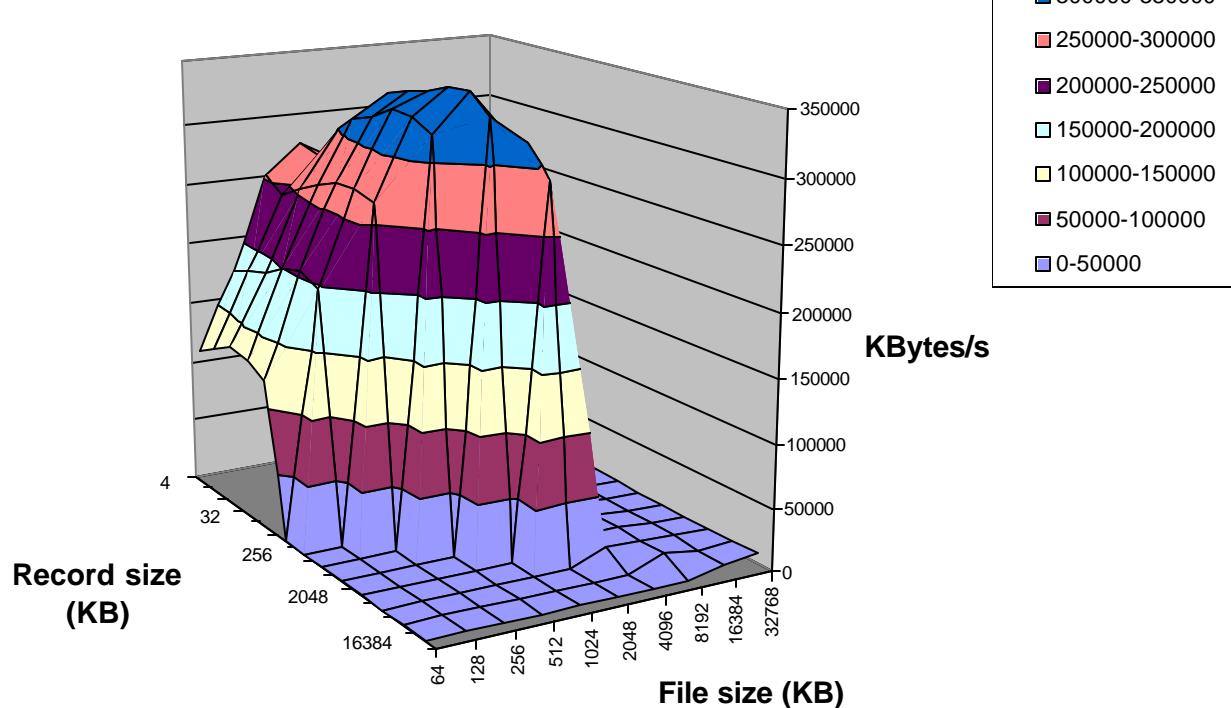
Diversity of operations

read / write / reread / rewrite

Influence of caches / system buffers

NFS latencies

## IPACS IO Write Benchmark



---

# Project Plan

Combine the goodies of PMB and IOZONE

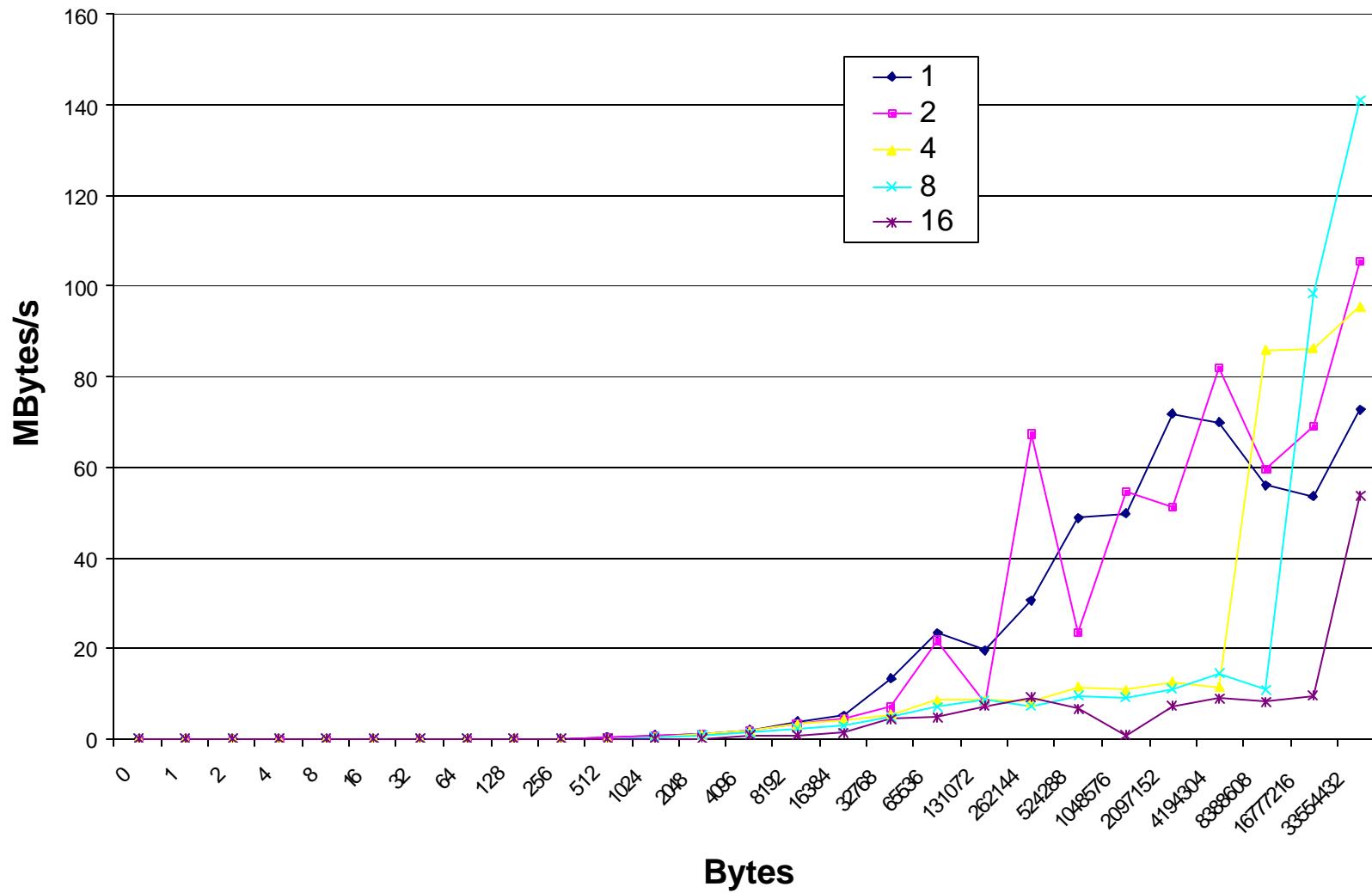
PMB prototype: PMB-IO

Measures basic IO characteristics with philosophy similar to message passing PMB

In particular, takes parallelism of IO into account

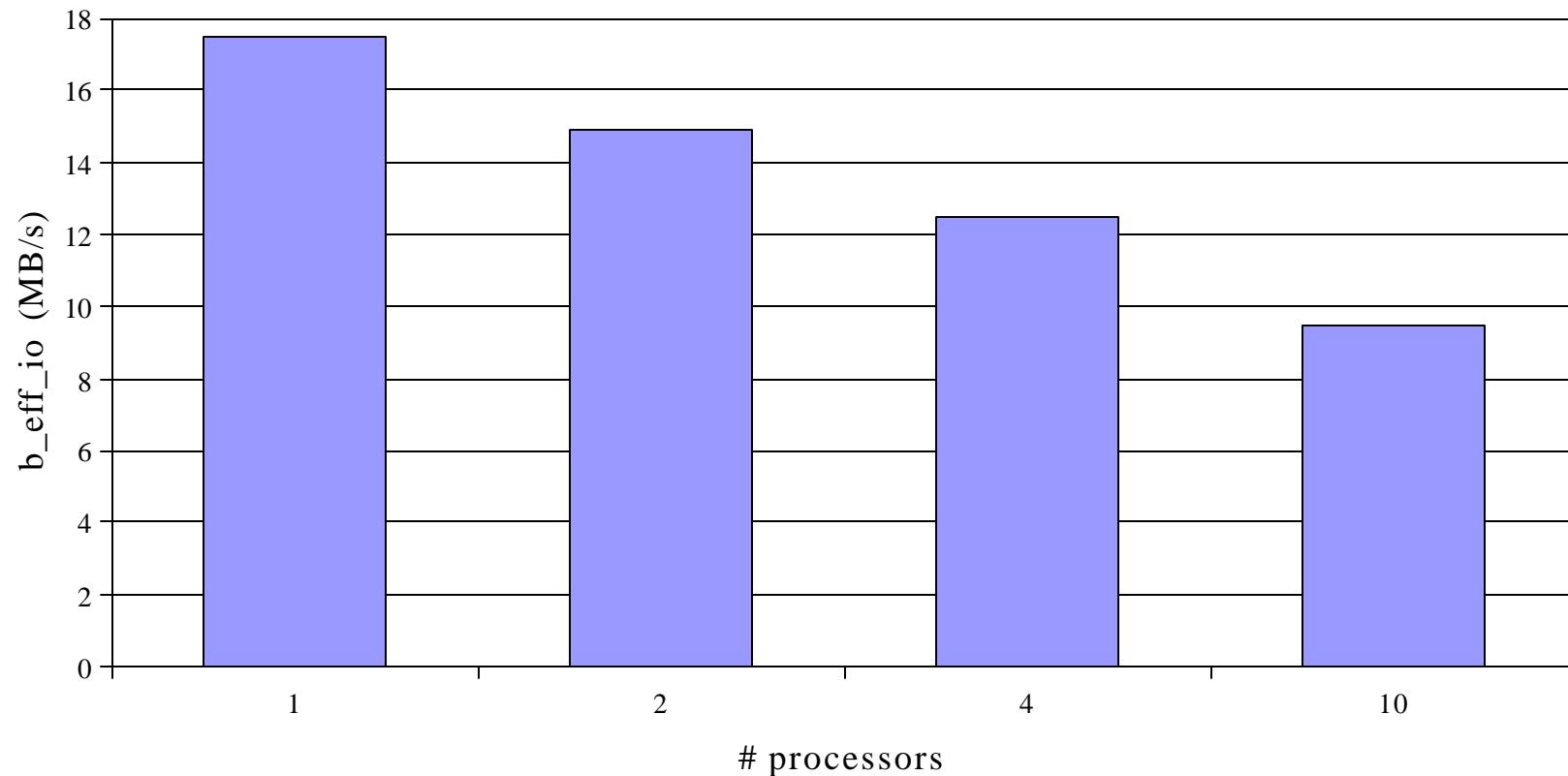
---

## PMB-IO IBM SP3: parallel write on 1-16 processors



# Cluster-I/O $b_{\text{eff\_io}}$

Measured on a Linux cluster with 5 SMP nodes



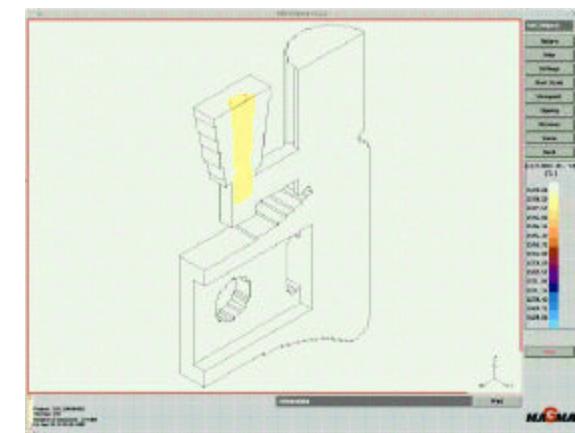
---

## Application benchmarks

---

# ParPac : Parallel Particel Code

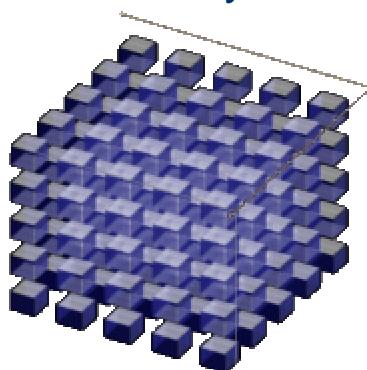
- solver for complex 3-dim flow problems
- developed at ITWM during the last 5 Years
- used in industry projects
- based on Lattice Boltzmann approach
- C++ Code , complex datastructures
- suited for very complex geometries
- fully parallelized
- automated domain decomposition
- dynamical load balance
- optimized communication paths



# Benchmark Example

Flow through a regular porous structure

- adapts easily to the machine size
- computes permeability
- ☒ compare results
- serial and parallel I/O
- ☒ measure restart process
- use different LBU granularities
- ☒ load balance, memory allocation



```
Konsole <8>
File Sessions Settings Help
=====
ParPac Application Benchmark
BGK Lattice Boltzmann simulation of a viscous fluid
streaming through a porous media
Fraunhofer ITWM - Germany
=====
Benchmark description:
Simulation of the dynamical behaviour of a      # * # * #
viscous fluid streaming through a regular      ----->
porous media under the influence of a      # * # * #
driving pressure gradient and calculation      ----->
of the permeability of the porous media.      # * # * #

System information:
Date      : Wed Dec 18 05:45:36 2002
Machine   : i686
System    : Linux
Release   : 2.4.19-lanl.18lnxi2
Version   : #3 SMP Wed Dec 11 15:49:00 MST 2002
Processors: 1536

Fluid parameters:
mass density      : 1.260e+03 kg/m^3
kinematic viscosity: 1.167e-03 m^2/s
pressure gradient : 7.000e+06 Pa/m

Benchmark parameters:
total no. of grid points      : 531441000
total no. of fluid points     : 465501736
material porosity           : 8.759e-01
size of each load balance unit: 1000

Benchmark results:
calculated permeability of the material: 1.992e-07 m^2

total simulation time       : 13.74 s
fluid grid point calculation rate: 338768320.85 1/s
floating point operation rate: 2.022e+02 Gflops

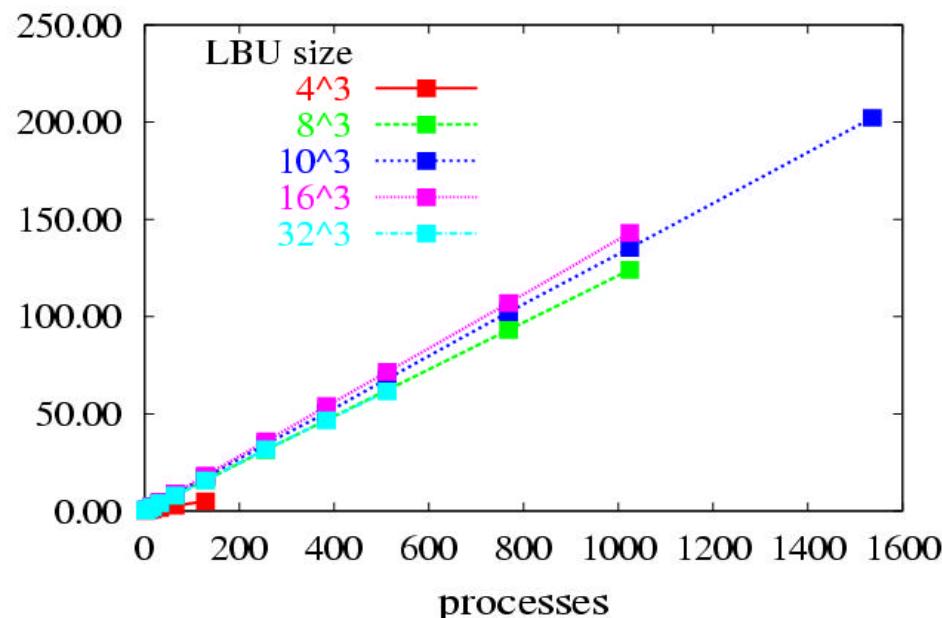
total I/O time for writing dump file: 0.00 s
I/O transfer rate to the disk    : 0.00 MBytes/s

End of benchmark
=====
```

# Parpacbench results

Compiler : gcc

ParPac Benchmark LANL Cluster (2 CPUs per Node)

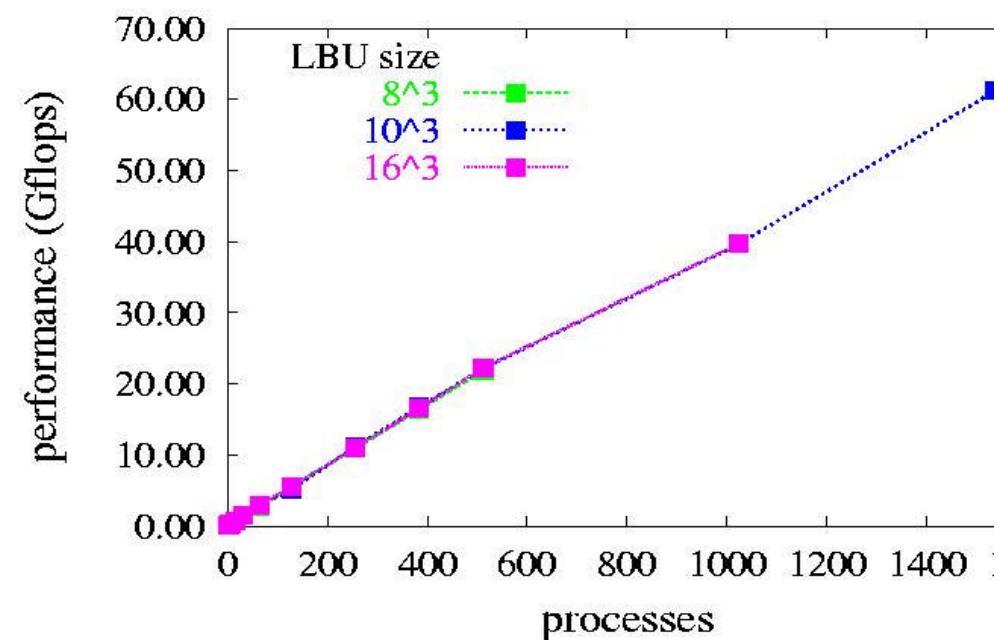


Percentage of Peak :

1536 CPU's : Peak : 7.37 TF  
200 GF **2.7 %**

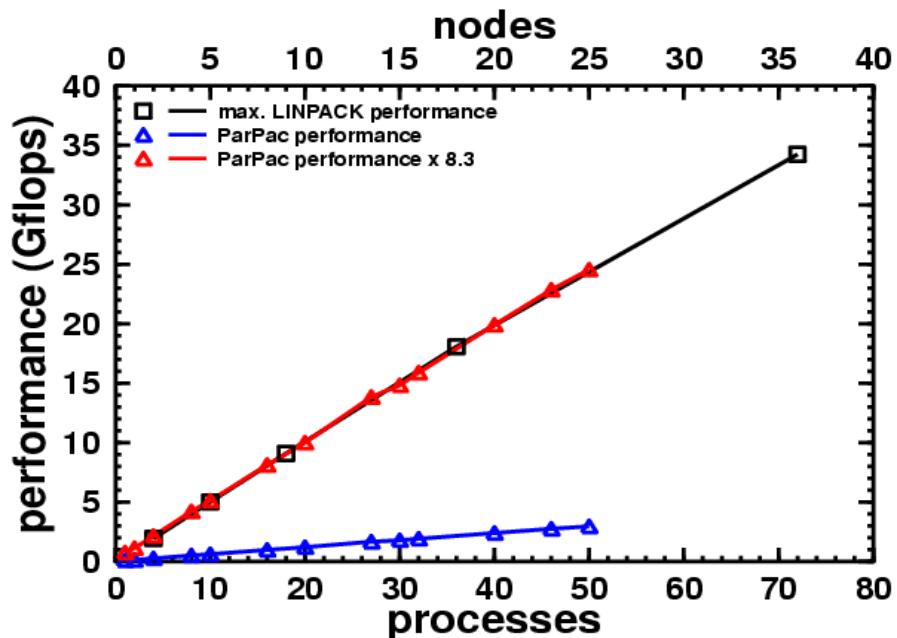
gcc + modified code (MPI)

ParPac Benchmark NERSC IBM SP



Peak : 2.3 TF  
61 GF **2.7 %**

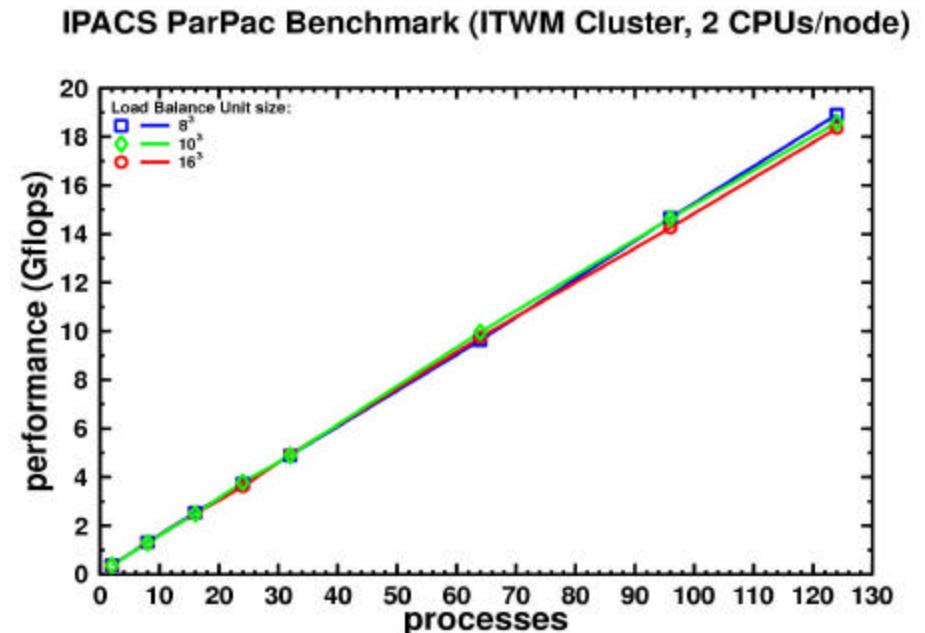
# Parpacbench results



Linpack = 8.3 \* Parpacbench

PC-Cluster-Myrinet

Pentium III 800 MHz



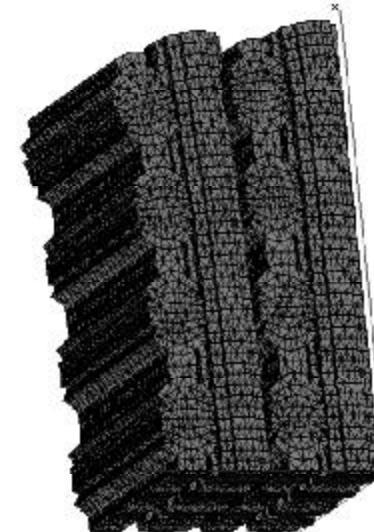
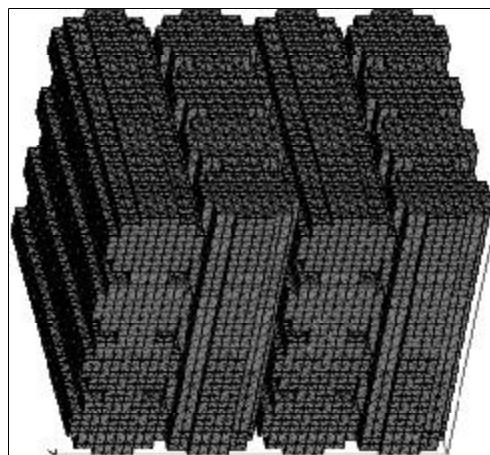
PC-Cluster P4 ( 2.4 GHz )

Myrinet 2000

# Linear Elasticity

Benchmark Code developed from scratch: ( C++ and PetSC)

Requirements : parallel geometry generation  
parallel Grid Generation  
scalability of benchmark example  
generate comparable results



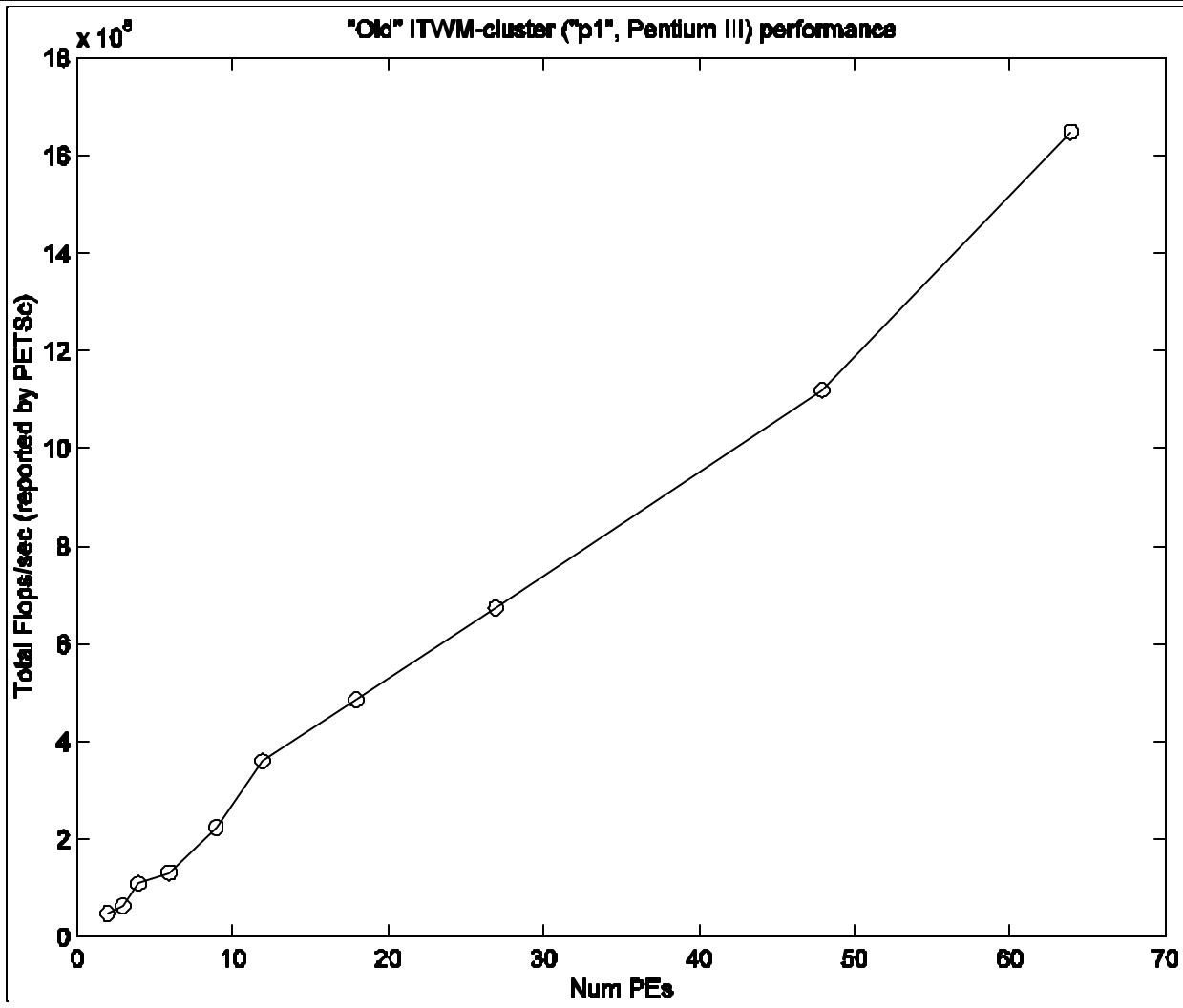
Fiber structure that is undergoing a compression

# First Results

PC –Cluster – Myrinet  
Pentium III 800 Mhz

Peak : 51.8 GF  
App : 1.7 GF  
3 %

**Measurements:**  
as reported by PetSC  
including:  
Geometry generation  
Grid generation



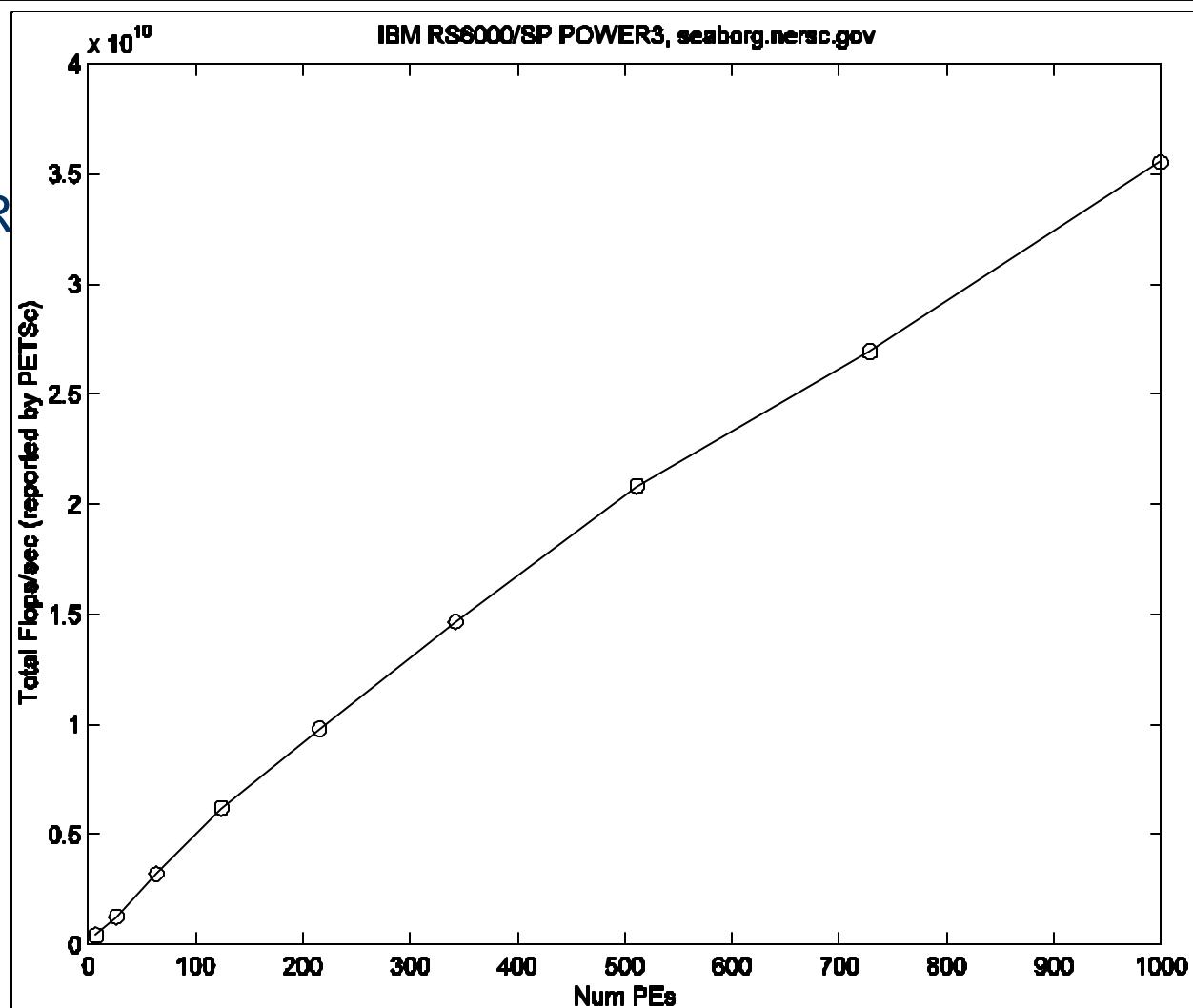
# First Results

IBM RS6000/SP POWER3 NER

Peak : 1.5 TF ( 1000 CPU)

App : 35 GF

2.3 %



---

## Benchmark environment

# Benchmark Client

Easy GUI development (Java)

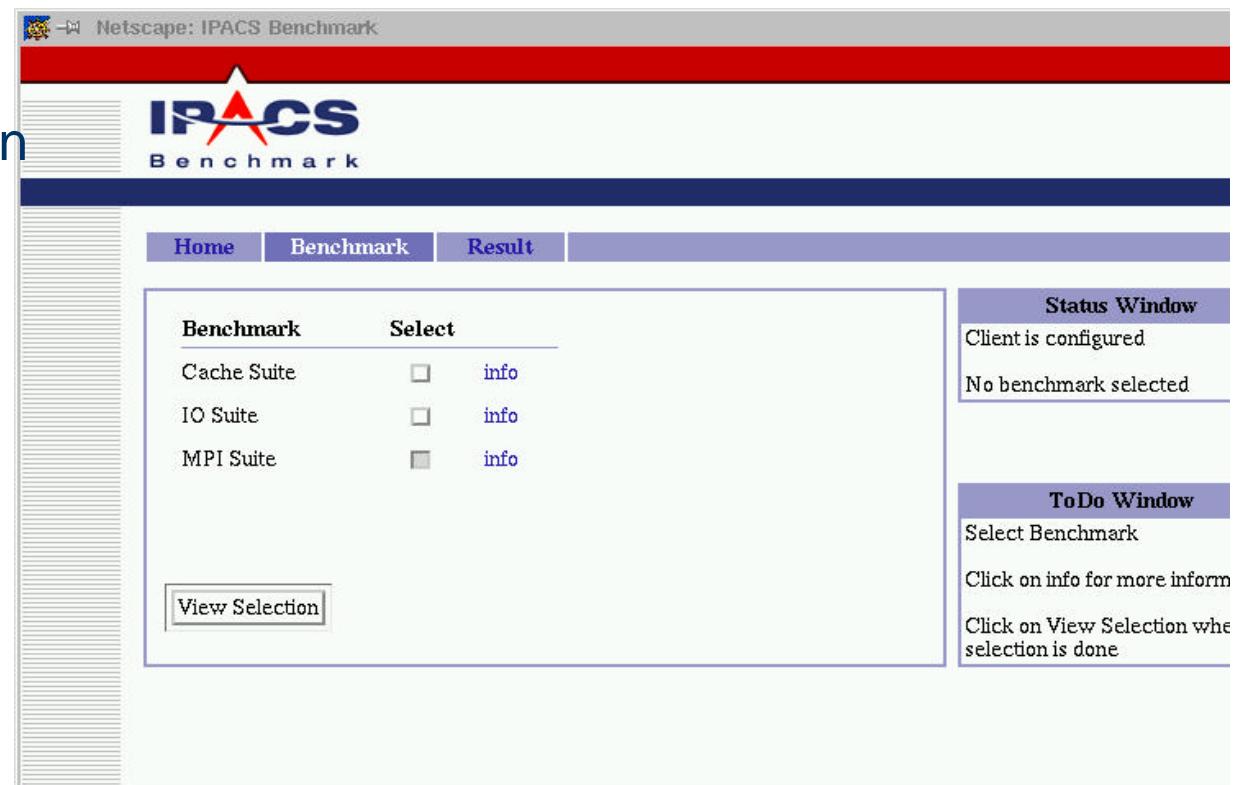
Detection of system configuration

Selecting and downloading benchmarks

Configuring benchmarks for HPC machines

Running benchmarks on HPC machines

Submitting results to IPACS server



# Benchmark Repository

Providing benchmark binaries  
and sources

Offering appropriate  
benchmarks for system  
configuration of HPC  
machines

Storing and presenting  
benchmark results

The screenshot shows a web page titled "Netscape: IPACS Benchmark Result". At the top is the IPACS Benchmark logo. Below it is a navigation bar with "Home", "Benchmark" (which is selected), and "Result". The main content area displays a table of benchmark results for the MPI Suite. The table has two columns: "Benchmark" and "Average MB/s". The data is as follows:

Benchmark	Average MB/s
<b>MPI Suite</b>	
PMB PingPong	75
PMB Sendrecv	30
PMB Allreduce	10

Below the table are three buttons: "Send to IPACS", "Save as XML", and "Save as HTML". To the right of the main content are two windows: "Status Window" showing "Result" and "ToDo Window" with the message "Details will show benchmark specific details" and "Send to IPACS will publish your results".



## Benchmark Results

Sponsored by:  
 Bundesministerium  
für Bildung  
und Forschung

Please select a benchmark to display

- Linpack
- ParPac
- Beff

Please select restrictions to create a list

All	Manufacturer
All	Installation Type
All	Processor Type
All	Computer Classification
All	Computer Family
All	Computer Type
All	Processor count
All	Processor count classification

And Select Fields To Display

- Manufacturer
- Computer name
- Processors
- Installation Type
- Installation Site
- Computer Type
- Year
- Computer Family

Show List    Reset



Fraunhofer Institut  
Techno- und  
Wirtschaftsmathematik

**IP**Acs  
Benchmark

---

# www.ipacs-benchmark.org

The screenshot shows a vintage web browser window titled "IPACS Benchmark - Netscape 6". The address bar displays the URL "http://www.ipacs-benchmark.org/". The main content area features a banner with the "IPACS Benchmark" logo, followed by a navigation menu with links to HOME, PROJECT, REPOSITORY, LINKS, INTERNAL, and CONTACT. Below the menu, there's a large button labeled "IPACS". To the right, a "Powered by" section includes the German Federal Ministry of Education and Research logo and the text "Bundesministerium für Bildung und Forschung". The central text on the page reads: "(Integrated Performance Analysis of Computer Systems) Benchmarks for distributed computer systems". It mentions funding from the Federal Department of Education, Science, Research and Technology (BMBF) through the 'High-Performance-Computing' program. The "Project Partners:" section lists several organizations: Fraunhofer Institut für Techno- und Wirtschaftsmathematik (ITWM), Pallas GmbH, T-Systems, Solutions for Research GmbH, Universität Mannheim, Rechenzentrum, and Universität Rostock, Technische Informatik. The "Project Synopsis and Goal:" section discusses the need for reliable benchmarks for PC Clusters and parallel systems, mentioning the Linpack performance used in TOP500. It also highlights cooperation with LBNL / NERSC and lists main work packages. The "Tools that support the user" section is partially visible at the bottom.

IPACS Benchmark - Netscape 6

http://www.ipacs-benchmark.org/

IPACS

Powered by:

Bundesministerium  
für Bildung  
und Forschung

(Integrated Performance Analysis of Computer Systems)

Benchmarks for distributed computer systems

The IPACS Project is funded by the Federal Department of Education, Science, Research and Technology (BMBF), in the program 'High-Performance-Computing'.

**Project Partners:**

- [Fraunhofer Institut für Techno- und Wirtschaftsmathematik \(ITWM\)](#)
- Pallas GmbH
- T-Systems, Solutions for Research GmbH
- Universität Mannheim, Rechenzentrum
- Universität Rostock, Technische Informatik

**Project Synopsis and Goal:**

With the continuing exponential growth of computer power and the growing relevance of PC Clusters system benchmarks, like NPB and others, have to be put on a new basis. At present we all miss reliable system benchmarks (other than the Linpack performance used in TOP500), which support the user in rating and obtaining computer systems - especially parallel systems.

In cooperation with colleagues at [LBNL / NERSC](#), IPACS wants to define a new basis for benchmarks measuring system performance of distributed systems. The following topics are the main work packages in IPACS:

- Development and propagation of scalable, portable and realistic benchmarks,
- Development of methods of measurement for parameters which are specific to applications and hardware,
- Prediction of performance of commercial codes and
- Tools that support the user.

The goal here is to develop methods of measurement based on new applications, low level, and grid benchmarks, which also allow performance prediction.

http://www.itwm.fhg.de/